ehp

LESSON:

Calculating Your Odds for Disease

Summary: Students use data from a case–control research study to calculate and interpret odds ratios and confidence intervals. Then they compare their results with the results reported by the researchers who conducted the study.

Lesson Type: Data Analysis—Students read and interpret data from tables.

***EHP* Article:** "Headliners: Polymorphisms Modify Breast Cancer Risk in Smokers"
EHP Student Edition, February 2007, p. A642
http://www.ehponline.org/docs/2006/114-11/niehsnews.html#poly

Objectives: By the end of this lesson, students should be able to

1. calculate odds ratios given case–control data in a table;
2. define and interpret confidence intervals; and
3. analyze and interpret odds ratio results in a case–control research study.

Class Time: 60–75 minutes

Grade Level: 11–12

Subjects Addressed: Biology, Environmental Sciences, General Science, Health

Prepping the Lesson (20 minutes)

INSTRUCTIONS:

1. Download the entire February 2007 *EHP Student Edition*, or download just the article "Headliners: Polymorphisms Modify Breast Cancer Risk in Smokers" at http://www.ehponline.org/docs/2006/114-11/niehsnews.html#poly.
2. Review the Background Information, Instructions, and Student Instructions.
3. Read the article "Headliners: Polymorphisms Modify Breast Cancer Risk in Smokers."
4. Make copies of the article and Student Instructions.
5. If your students have very little background knowledge on cancer, genetics, and the relationship between the two, you may need to prepare a brief introduction on the topic or plan to implement this lesson during a genetics unit. Alternatively, you could exclude Step 7b, which asks students to reach conclusions about genetic polymorphisms and breast cancer rates.

MATERIALS (per student):

- 1 copy of *EHP Student Edition*, February 2007, or 1 copy of the article "Headliners: Polymorphisms Modify Breast Cancer Risk in Smokers"
- 1 copy of the Student Instructions
- Calculator

VOCABULARY:

- case
- case–control study
- confidence interval
- control
- genotype
- nucleotide
- nucleotide excision repair (NER) gene
- odds ratio
- peripheral blood lymphocyte
- polymorphism
- single-nucleotide polymorphism (SNP)

BACKGROUND INFORMATION:

This lesson uses data from a published case–control research study looking at the relationships between polymorphisms in nucleotide excision repair (NER) genes, exposure to cigarette smoke, and breast cancer (Mechanic, et al. 2006). Selected data are taken from the study and used by the students to calculate odds ratios, a way to show relationships in research studies of this type. Although actual data from the study are used by the students, the odds ratios calculated by the students are not the same as those reported in the study. The study used advanced statistical techniques to adjust for confounding factors between cases and controls, which could not be duplicated by the students. The students' calculations do, however, illustrate the principles involved and produce similar results to those reported in the actual study. Additional background information about the study is provided in the Student Instructions.

Reference:

Mechanic LE, Millikan RC, Player J, de Cotret AR, Winkel S, Worley K, et al. 2006. Polymorphisms in nucleotide excision repair genes, smoking and breast cancer in African Americans and whites: a population-based case–control study. Carcinogenesis 27:1377–1385.

RESOURCES:

Environmental Health Perspectives, Environews by Topic page, http://ehp.niehs.nih.gov/. Choose Genetic Research, Research Issues and Initiatives, Risk Assessment

Bland JM, Altman DG. 2000. The odds ratio. BMJ 320:1468; available at: http://www.bmj.com/cgi/content/full/320/7247/1468

National Center for Biotechnology Information, SNPs: variations on a theme, http://www.ncbi.nlm.nih.gov/About/priimer/snps.html

National Human Genome Research Institute homepage, http://www.genome.gov/

U.S. Department of Energy Office of Science Human Genome Program homepage, http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml

Wikipedia, Odds ratio, http://en.wikipedia.org/wiki/Odds_ratio

Implementing the Lesson

INSTRUCTIONS:

1. Ask students what they think is meant by an odds ratio. It may be helpful to have students define "odds" and "ratio" independently. The term "odds" refers to the probability or likelihood of an event happening, and a "ratio" is simply comparing two numbers. The term "odds ratio" means a way to determine if the probability of a specific event is the same for two different groups.
2. Tell the students they are going to use real data from a research study to calculate whether there are differences between two groups of women (white and black) with respect to their risk for breast cancer. Tell the students they will also compare smokers to nonsmokers as well as genetic factors as they relate to race and breast cancer risk.
3. Have students complete the Student Instructions, working either individually or in small groups.
4. Lead a class discussion about their results.

NOTES & HELPFUL HINTS:

1. The entire lesson could be done as homework.
2. The lesson could be expanded by having the students look at the results and tables in the published case–control research study.

Aligning with Standards

SKILLS USED OR DEVELOPED:

- Communication (note-taking—oral, written)
- Comprehension (listening, reading)
- Computation
- Critical thinking and response
- Research
- Tables and figures (creating, reading)

SPECIFIC CONTENT ADDRESSED:

- Environmental health
- Odds ratios
- Case–control studies

- Confidence intervals
- Genetic basis of disease
- Breast cancer
- Smoking, risk

NATIONAL SCIENCE EDUCATION STANDARDS MET:

Science Content Standards

Unifying Concepts and Processes Standard
- Systems, order, and organization
- Evidence, models, and explanation
- Change, constancy, and measurement
- Form and function

Science as Inquiry Standard
- Abilities necessary to do scientific inquiry
- Understanding about scientific inquiry

Physical Science Standard
- Structure and properties of matter
- Chemical reactions

Life Science Standard
- The cell
- Molecular basis of heredity
- Interdependence of organisms
- Behavior of organisms

Science in Personal and Social Perspectives Standard
- Personal and community health
- Environmental quality
- Natural and human-induced hazards
- Science and technology in local, national, and global challenges

History and Nature of Science Standard
- Science as a human endeavor
- Nature of scientific knowledge
- Historical perspectives

Assessing the Lesson

Step 2: A case–control research study was done to examine the relationships between polymorphisms (or several possible expressions of a gene) in nucleotide excision repair (NER) genes, smoking, and breast cancer in black and white women. Smoking and exposure to passive smoke are known risk factors for a number of different types of cancer. Cigarette smoke is believed to cause genetic mutations that lead to the increased risk of cancer. You will be introduced to several genetics terms related to this activity, including "NER genes," "polymorphisms," and "single-nucleotide polymorphisms (SNPs)."

- NER genes are responsible for repairing damaged DNA by cutting out the damaged section and replacing it with an undamaged strand.
- Polymorphisms are genetic differences in DNA sequences between individuals that may underlie differences in health.
- SNPs are changes to one single nucleotide in a person's genetic sequence—for example, let's say a DNA segment has the nucleotide sequence of AAAACGAGCA, and the first A (adenine) is replaced with a T (thymine). The changed sequence now reads TAAACGAGCA. Depending on where a SNP occurs in the genetic sequence, there may be no effect or there could be a significant change, such as affecting the ability of a protein to do its job.
- A number of SNPs have been discovered in NER genes ("at-risk" genotypes) that may reduce their repair capabilities and therefore increase the risk for cancer.

This case–control research study was initiated to answer the questions of 1) whether black and white women who have breast cancer are more likely to be exposed to cigarette smoke; and 2) whether black and white women who have breast cancer are more likely to be exposed to cigarette smoke and have "at-risk" genotypes. What do you predict would be answer to these two questions? Explain why.

Answers will vary. A typical answer might be yes to both questions for both black and white women. If cigarette smoke is a known human carcinogen, increased exposure to cigarette smoke might be expected to be associated with higher numbers of breast cancer in women. If DNA repair mechanisms are also impaired, and cigarette smoke causes cancer by causing mutations, women who smoke and have "at-risk" genotypes might be expected to be associated with greater risk of breast cancer.

Step 3: Tables 1 and 2 show the number of black and white women and smoking exposures from the case–control research study.

a. Complete the tables, calculating the odds and odds ratios that resulted from the comparisons in each table.

Table 1
Odds Ratio of Black Women with Breast Cancer Who Smoked to Black Women without Breast Cancer Who Smoked

	Not exposed to active and passive smoke (Column 1)	**Active smoker for greater than 20 years** (Column 2)	**Odds for being an active smoker** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds of cases who were active smokers divided by odds of controls who were active smokers)
# of Cases (Women have breast cancer)	161	198	1.230	1.60
# of Controls (Women do not have breast cancer)	172	132	0.767	

Table 2
Odds Ratio of White Women with Breast Cancer Who Smoked to White Women without Breast Cancer Who Smoked

	Not exposed to active and passive smoke (Column 1)	**Active smoker for greater than 20 years** (Column 2)	**Odds for being an active smoker** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds of cases who were active smokers divided by odds of controls who were active smokers)
# of Cases (Women have breast cancer)	287	335	1.167	0.94
# of Controls (Women do not have breast cancer)	235	292	1.243	

b. What do these results indicate about whether black and white women who have breast cancer are more likely to be exposed to cigarette smoke?

Black women who have breast cancer are more likely to be active smokers and therefore, more likely to be exposed to cigarette smoke since the odds ratio in Table 1 is greater than 1. This is not true for white women who smoke, since the odds ratio in Table 2 is less than 1.

Step 4: Tables 3 and 4 show the number of black and white women with four or more "at-risk" genotypes and their smoking exposures from the case–control research study.

a. Complete the tables calculating the odds and odds ratios that resulted from the comparisons in each table.

Table 3
Odds Ratio of Black Women with Breast Cancer Who Smoked and Have 4 or More "At-Risk" Genotypes to Black Women Without Breast Cancer Who Smoked and Have 4 or More "At-Risk" Genotypes

	Not exposed to active and passive smoke and had 4 or more "at-risk" genotypes (Column 1)	**Active smoker for greater than 20 years and had 4 or more "at-risk" genotypes** (Column 2)	**Odds for being an active smoker and having 4 or more "at-risk" genotypes** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds of cases who were active smokers and had 4 or more "at-risk" genotypes divided by odds of controls who were active smokers and had 4 or more "at-risk" genotypes)
# of Cases (Women have breast cancer)	35	51	1.457	3.31
# of Controls (Women do not have breast cancer)	50	22	0.440	

Table 4
Odds Ratio of White Women with Breast Cancer Who Smoked and Have 4 or More "At-Risk" Genotypes to White Women Without Breast Cancer Who Smoked and Have 4 or More "At-Risk" Genotypes

	Not exposed to active and passive smoke and had 4 or more "at-risk" genotypes (Column 1)	**Active smoker for greater than 20 years and had 4 or more "at-risk" genotypes** (Column 2)	**Odds for being an active smoker and having 4 or more "at-risk" genotypes** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds of cases who were active smokers and had 4 or more "at-risk" genotypes divided by odds of controls who were active smokers and had 4 or more "at-risk" genotypes)
# of Cases (Women have breast cancer)	63	75	1.190	1.05
# of Controls (Women do not have breast cancer)	59	67	1.136	

b. What do these results indicate about whether black and white women who have breast cancer are more likely to be exposed to cigarette smoke and have "at-risk" genotypes?

Black women who have breast cancer are more likely to be active smokers and have four or more "at-risk" genotypes since the odds ratio in Table 1 is greater than 1. Since the odds ratio in Table 3 is greater than the odds ratio in Table 1, the additional risk factor of "at-risk" genotypes seems to increase the chances of breast cancer and smoking being associated. White women who have breast cancer are more likely to smoke and have four or more "at-risk" genotypes since the odds ratio in Table 4 is slightly greater than 1.

c. Research studies don't report odds ratios as precise numbers, but qualify these numbers by giving confidence intervals. Confidence intervals are a range of numbers indicating the possible range of values that contain the true value with 95% probability. According to the original research paper, the confidence interval for the odds ratio in Table 4 is 0.6 to 1.7. This means the true odds ratio lies between 0.6 and 1.7 with 95% probability. Although the odds ratio is reported as 1.05, there is a 95% probability that it could be as low as 0.6 or as high as 1.7. Given this confidence interval, what can you conclude about whether white women with breast cancer are more likely to be exposed to cigarette smoke and have "at-risk" genotypes? (Consider what <1, =1, and >1 odds ratios mean.)

Based on the confidence interval, the odds ratio could be less than 1, equal to 1 or above 1. Therefore, one can not say that white women who have breast cancer are associated with smoking and have 'at risk' genotypes.

Step 5: How do your predictions from Step 2 compare with the results of the case–control research study for which you just calculated the odds ratios? Do the data support or not support your hypothesis? Why or why not?

Answers will vary. Look for consistency between what students say they said and their conclusions. Be sure they provide clear, logical explanations.

Step 6: Read the article "Headliners: Polymorphisms Modify Breast Cancer Risk in Smokers." This is a summary of the actual research study published in the journal *Carcinogenesis*. Compare the conclusions you reached from looking at some of the data from the research study and the researchers' conclusions.

Answers will vary. The article found smoking was more strongly associated with breast cancer in black women than in white women. The association increased even more for black women with particular patterns of polymorphisms when combined with different smoking characteristics including amount of smoking, duration, time since smoking cessation, age at smoking initiation, and being a former smoker.

Step 7: a. Using the information in the "Headliners" article and your own ideas, list two possible genetic or nongenetic reasons for the different results in the case–control research study between black and white women.

Answers will vary. Encourage the students to be as specific as possible in their answers. For example, just stating "there may be genetic differences" is not sufficient. The authors of the study provide the following explanations as to why black women with breast cancer are associated with higher exposures to cigarette smoke while white women are not. Some of these explanations are more complex or sophisticated than the answers students may provide, but these do provide some guidelines on which to gauge student answers.

- Black women may be exposed to higher amounts of secondhand tobacco smoke than white women, increasing their risk of breast cancer.
- Genetically based hormonal differences may exist between black and white women, increasing the risk of breast cancer in black women.
- Adjustments for confounding factors in the study such as income, education, lactation, body mass index, and use of birth control pills and hormone replacement therapy may not be accurate, resulting in a greater number of breast cancer cases for black women. (Students likely do not know that confounding factors may be adjusted for statistically, but they may provide any one of these factors, such as income differences, as a possible contributing factor. This is fine.)
- The number of participants in the different categories was small, resulting in wide confidence intervals for some odds ratios. Therefore, the observed greater number of cases for black women who smoked and had "at-risk" genotypes could have resulted by chance.
- Black women may have other unidentified susceptibility factors that increased their risk of breast cancer.

b. As mentioned previously, the article discusses genetic polymorphisms, which means there are several viable expressions for a single gene resulting in different phenotypes or physical expressions of that gene (for example, blue eyes versus brown eyes). NER genes were specifically investigated in this research. NER genes play a key role in repairing damage to DNA. Why might differences in NER genes in black and white women impact breast cancer rates? Think about how DNA damage and cancer relate to each other, as well as your understanding of how genes function.

   Student answers may vary but should address some of the following points: if DNA is damaged and not repaired, the risk for cancer is higher; black women may have NER genes that are less effective at repairing DNA or require the presence of specific chemicals that may be missing; the NER genes in black women may be more easily "down-regulated," or turned off by environmental factors.

Authors and Reviewers

Authors: Barry Schlegel and Laura Hemminger, University of Medicine and Dentistry of New Jersey–School of Public Health
Reviewers: Stephanie Bishop, Susan Booker, Erin Dooley, Stefani Hines, and Joseph Tart

Give us your feedback! Send comments about this lesson to ehpscienceed@niehs.nih.gov.

ehp

Student Instructions: Calculating Your Odds for Disease

Step 1: Health scientists sometimes use case–control studies to identify factors that are associated with increased risk of disease. Individuals with the disease (cases) are compared to individuals without the disease (controls). Controls are selected and matched to the cases so that the two groups have similar characteristics (such as age). Cases and controls differ only by the factor(s) believed to be connected to the disease.

Health scientists sometimes use odds ratios to show the relationship between disease and risk factors between the two groups. Odds are a way to express the probability of an event in one group. The odds ratio is the odds or chances of something happening in the group of cases divided by the odds or chances of something happening in the group of controls. An odds ratio equal to 1 means the groups have equal probabilities of the event. An odds ratio greater than 1 means cases have a greater probability of the event, and an odds ratio less than 1 means the cases have less probability of the event.

As an example, let's say you selected as your cases 100 women with breast cancer. Of these 100 women, 20 smoked cigarettes. The proportion of women who smoked is 0.2 (20/100, or 20%). The proportion of women who didn't smoke is 0.8 (80/100, or 80%). For this group of women with breast cancer, the odds of being a smoker versus not being a smoker is 0.25 (0.2 divided by 0.8, or 25%).

Another 100 women without breast cancer were selected as controls. Of these 100 women, 10 smoked cigarettes. The proportion of women who smoked is 0.1 (10/100, or 10%). The proportion of women who didn't smoke is 0.9 (90/100, or 90%). For this group of women without breast cancer, the odds of being a smoker versus not being a smoker is 0.11 (0.1 divided by 0.9 or 11%).

The odds ratio between the cases and the controls in this example would be 2.27 (0.25 divided by 0.11). Since the odds ratio is greater than 1, women with breast cancer are more likely to smoke than women without breast cancer. Smoking seems to be associated with breast cancer.

$$\text{Odds Ratio} = \frac{\text{Odds of a woman with breast cancer being a smoker}}{\text{Odds of a woman without breast cancer being a smoker}} = \frac{0.25}{0.11} = 2.27$$

This example is pictured in the following table showing the number of cases and controls who smoked and didn't smoke and how the odds ratio can be calculated from the numbers in the cells of the table:

Odds Ratio of Women with Breast Cancer Who Smoked to Women without Breast Cancer Who Smoked				
	Did not smoke (Column 1)	**Smoked** (Column 2)	**Odds for being an smoker** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds for cases who smoked divided by odds for controls who smoked)
# of Cases (Women have breast cancer)	80	20	0.25 (20/80)	2.27 (0.25/0.11)
# of Controls (Women do not have breast cancer)	90	10	0.11 (10/90)	

National Institute of Environmental Health Sciences
Community Outreach & Education Program

Step 2: A case–control research study was done to examine the relationships between polymorphisms (or several possible expressions of a gene) in nucleotide excision repair (NER) genes, smoking, and breast cancer in black and white women. Smoking and exposure to passive smoke are known risk factors for a number of different types of cancer. Cigarette smoke is believed to cause genetic mutations that lead to the increased risk of cancer. You will be introduced to several genetics terms related to this activity, including "NER genes," "polymorphisms," and "single-nucleotide polymorphisms (SNPs)."

- NER genes are responsible for repairing damaged DNA by cutting out the damaged section and replacing it with an undamaged strand.
- Polymorphisms are genetic differences in DNA sequences between individuals that may underlie differences in health.
- SNPs are changes to one single nucleotide in a person's genetic sequence—for example, let's say a DNA segment has the nucleotide sequence of AAAACGAGCA, and the first A (adenine) is replaced with a T (thymine). The changed sequence now reads TAAACGAGCA. Depending on where a SNP occurs in the genetic sequence, there may be no effect or there could be a significant change, such as affecting the ability of a protein to do its job.
- A number of SNPs have been discovered in NER genes ("at-risk" genotypes) that may reduce their repair capabilities and therefore increase the risk for cancer.

This case–control research study was initiated to answer the questions of 1) whether black and white women who have breast cancer are more likely to be exposed to cigarette smoke; and 2) whether black and white women who have breast cancer are more likely to be exposed to cigarette smoke and have "at-risk" genotypes. What do you predict would be answer to these two questions? Explain why.

Step 3: Tables 1 and 2 show the number of black and white women and smoking exposures from the case–control research study.

a. Complete the tables, calculating the odds and odds ratios that resulted from the comparisons in each table.

Table 1
Odds Ratio of Black Women with Breast Cancer Who Smoked to Black Women without Breast Cancer Who Smoked

	Not exposed to active and passive smoke (Column 1)	**Active smoker for greater than 20 years** (Column 2)	**Odds for being an active smoker** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds of cases who were active smokers divided by odds of controls who were active smokers)
# of Cases (Women have breast cancer)	161	198		
# of Controls (Women do not have breast cancer)	172	132		

Table 2
Odds Ratio of White Women with Breast Cancer Who Smoked to White Women without Breast Cancer Who Smoked

	Not exposed to active and passive smoke (Column 1)	**Active smoker for greater than 20 years** (Column 2)	**Odds for being an active smoker** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds of cases who were active smokers divided by odds of controls who were active smokers)
# of Cases (Women have breast cancer)	287	335		
# of Controls (Women do not have breast cancer)	235	292		

b. What do these results indicate about whether black and white women who have breast cancer are more likely to be exposed to cigarette smoke?

Step 4: Tables 3 and 4 show the number of black and white women with four or more "at-risk" genotypes and their smoking exposures from the case–control research study.

a. Complete the tables calculating the odds and odds ratios that resulted from the comparisons in each table.

Table 3
Odds Ratio of Black Women with Breast Cancer Who Smoked and Have 4 or More "At-Risk" Genotypes to Black Women without Breast Cancer Who Smoked and Have 4 or More "At-Risk" Genotypes

	Not exposed to active and passive smoke and had 4 or more "at-risk" genotypes (Column 1)	**Active smoker for greater than 20 years and had 4 or more "at-risk" genotypes** (Column 2)	**Odds for being an active smoker and having 4 or more "at-risk" genotypes** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds of cases who were active smokers and had 4 or more "at-risk" genotypes divided by odds of controls who were active smokers and had 4 or more "at-risk" genotypes)
# of Cases (Women have breast cancer)	35	51		
# of Controls (Women do not have breast cancer)	50	22		

Table 4
Odds Ratio of White Women with Breast Cancer Who Smokeed and Have 4 or More "At-Risk" Genotypes to White Women without Breast Cancer Who Smoked and Have 4 or More "At-Risk" Genotypes

	Not exposed to active and passive smoke and had 4 or more "at-risk" genotypes (Column 1)	**Active smoker for greater than 20 years and had 4 or more "at-risk" genotypes** (Column 2)	**Odds for being an active smoker and having 4 or more "at-risk" genotypes** (# from Column 2 divided by # from Column 1)	**Odds Ratio** (Odds of cases who were active smokers and had 4 or more "at-risk" genotypes divided by odds of controls who were active smokers and had 4 or more "at-risk" genotypes)
# of Cases (Women have breast cancer)	63	75		
# of Controls (Women do not have breast cancer)	59	67		

b. What do these results indicate about whether black and white women who have breast cancer are more likely to be exposed to cigarette smoke and have "at-risk" genotypes?

c. Research studies don't report odds ratios as precise numbers, but qualify these numbers by giving confidence intervals. Confidence intervalss are a range of numbers indicating the possible range of values that contain the true value with 95% probability. According to the original research paper, the confidence interval for the odds ratio in Table 4 is 0.6 to 1.7. This means the true odds ratio lies between 0.6 and 1.7 with 95% probability. Although the odds ratio is reported as 1.05, there is a 95% probability that it could be as low as 0.6 or as high as 1.7. Given this confidence interval, what can you conclude about whether white women with breast cancer are more likely to be exposed to cigarette smoke and have "at-risk" genotypes? (Consider what <1, =1, and >1 odds ratios mean.)

Step 5: How do your predictions from Step 2 compare with the results of the case–control research study for which you just calculated the odds ratios? Do the data support or not support your hypothesis? Why or why not?

Step 6: Read the article "Headliners: Polymorphisms Modify Breast Cancer Risk in Smokers." This is a summary of the actual research study published in the journal *Carcinogenesis*. Compare the conclusions you reached from looking at some of the data from the research study and the researchers' conclusions.

Step 7: a. Using the information in the "Headliners" article and your own ideas, list two possible genetic or nongenetic reasons for the different results in the case–control research study between black and white women.

b. As mentioned previously, the article discusses genetic polymorphisms, which means there are several viable expressions for a single gene resulting in different phenotypes or physical expressions of that gene (for example, blue eyes versus brown eyes). NER genes were specifically investigated in this research. NER genes play a key role in repairing damage to DNA. Why might differences in NER genes in black and white women impact breast cancer rates? Think about how DNA damage and cancer relate to each other, as well as your understanding of how genes function.